

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/558,149A
Source: TFW16
Date Processed by STIC: 03/14/2006

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/09/558,149A

TIME: 11:52:44

Input Set : A:\980232_1.TXT

Output Set: N:\CRF4\03142006\I558149A.raw

```

4 <110> APPLICANT: NICOLAIDES, NICHOLAS
5     VOGELSTEIN, BERT
6     KINZLER, KINZLER
8 <120> TITLE OF INVENTION: A METHOD FOR GENERATING HYPERMUTABLE
9     ORGANISMS
11 <130> FILE REFERENCE: 01107.000004
13 <140> CURRENT APPLICATION NUMBER: 09/558149A
14 <141> CURRENT FILING DATE: 2000-04-26
16 <150> PRIOR APPLICATION NUMBER: 09/059461
17 <151> PRIOR FILING DATE: 1998-04-14
19 <160> NUMBER OF SEQ ID NOS: 8
21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 2771
25 <212> TYPE: DNA
26 <213> ORGANISM: Homo sapiens
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (25)...(2610)
32 <400> SEQUENCE: 1
33 cgaggcggat cgggtgttgc atcc atg gag cga gct gag agc tcg agt aca      51
34                               Met Glu Arg Ala Glu Ser Ser Ser Thr
35                               1           5
37 gaa cct gct aag gcc atc aaa cct att gat cgg aag tca gtc cat cag      99
38 Glu Pro Ala Lys Ala Ile Lys Pro Ile Asp Arg Lys Ser Val His Gln
39 10           15           20           25
41 att tgc tct ggg cag gtg gta ctg agt cta agc act gcg gta aag gag      147
42 Ile Cys Ser Gly Gln Val Val Leu Ser Leu Ser Thr Ala Val Lys Glu
43           30           35           40
45 tta gta gaa aac agt ctg gat gct ggt gcc act aat att gat cta aag      195
46 Leu Val Glu Asn Ser Leu Asp Ala Gly Ala Thr Asn Ile Asp Leu Lys
47           45           50           55
49 ctt aag gac tat gga gtg gat ctt att gaa gtt tca gac aat gga tgt      243
50 Leu Lys Asp Tyr Gly Val Asp Leu Ile Glu Val Ser Asp Asn Gly Cys
51           60           65           70
53 ggg gta gaa gaa gaa aac ttc gaa ggc tta act ctg aaa cat cac aca      291
54 Gly Val Glu Glu Glu Asn Phe Glu Gly Leu Thr Leu Lys His His Thr
55           75           80           85
57 tct aag att caa gag ttt gcc gac cta act cag gtt gaa act ttt ggc      339
58 Ser Lys Ile Gln Glu Phe Ala Asp Leu Thr Gln Val Glu Thr Phe Gly
59 90           95           100           105
61 ttt cgg ggg gaa gct ctg agc tca ctt tgt gca ctg agc gat gtc acc      387
62 Phe Arg Gly Glu Ala Leu Ser Ser Leu Cys Ala Leu Ser Asp Val Thr

```

RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/09/558,149A

TIME: 11:52:44

Input Set : A:\980232_1.TXT

Output Set: N:\CRF4\03142006\I558149A.raw

63		110		115		120	
65	att tct acc tgc cac gca tgc gcg aag gtt gga act cga ctg atg ttt					435	
66	Ile Ser Thr Cys His Ala Ser Ala Lys Val Gly Thr Arg Leu Met Phe						
67		125		130		135	
69	gat cac aat ggg aaa att atc cag aaa acc ccc tac ccc cgc ccc aga					483	
70	Asp His Asn Gly Lys Ile Ile Gln Lys Thr Pro Tyr Pro Arg Pro Arg						
71		140		145		150	
73	ggg acc aca gtc agc gtg cag cag tta ttt tcc aca cta cct gtg cgc					531	
74	Gly Thr Thr Val Ser Val Gln Gln Leu Phe Ser Thr Leu Pro Val Arg						
75		155		160		165	
77	cat aag gaa ttt caa agg aat att aag aag gag tat gcc aaa atg gtc					579	
78	His Lys Glu Phe Gln Arg Asn Ile Lys Lys Glu Tyr Ala Lys Met Val						
79	170		175		180		185
81	cag gtc tta cat gca tac tgt atc att tca gca ggc atc cgt gta agt					627	
82	Gln Val Leu His Ala Tyr Cys Ile Ile Ser Ala Gly Ile Arg Val Ser						
83		190		195		200	
85	tgc acc aat cag ctt gga caa gga aaa cga cag cct gtg gta tgc aca					675	
86	Cys Thr Asn Gln Leu Gly Gln Gly Lys Arg Gln Pro Val Val Cys Thr						
87		205		210		215	
89	ggt gga agc ccc agc ata aag gaa aat atc ggc tct gtg ttt ggg cag					723	
90	Gly Gly Ser Pro Ser Ile Lys Glu Asn Ile Gly Ser Val Phe Gly Gln						
91		220		225		230	
93	aag cag ttg caa agc ctc att cct ttt gtt cag ctg ccc cct agt gac					771	
94	Lys Gln Leu Gln Ser Leu Ile Pro Phe Val Gln Leu Pro Pro Ser Asp						
95		235		240		245	
97	tcc gtg tgt gaa gag tac ggt ttg agc tgt tgc gat gct ctg cat aat					819	
98	Ser Val Cys Glu Glu Tyr Gly Leu Ser Cys Ser Asp Ala Leu His Asn						
99	250		255		260		265
101	ctt ttt tac atc tca ggt ttc att tca caa tgc acg cat gga gtt gga					867	
102	Leu Phe Tyr Ile Ser Gly Phe Ile Ser Gln Cys Thr His Gly Val Gly						
103		270		275		280	
105	agg agt tca aca gac aga cag ttt ttc ttt atc aac cgg cgg cct tgt					915	
106	Arg Ser Ser Thr Asp Arg Gln Phe Phe Phe Ile Asn Arg Arg Pro Cys						
107		285		290		295	
109	gac cca gca aag gtc tgc aga ctc gtg aat gag gtc tac cac atg tat					963	
110	Asp Pro Ala Lys Val Cys Arg Leu Val Asn Glu Val Tyr His Met Tyr						
111		300		305		310	
113	aat cga cac cag tat cca ttt gtt gtt ctt aac att tct gtt gat tca					1011	
114	Asn Arg His Gln Tyr Pro Phe Val Val Leu Asn Ile Ser Val Asp Ser						
115		315		320		325	
117	gaa tgc gtt gat atc aat gtt act cca gat aaa agg caa att ttg cta					1059	
118	Glu Cys Val Asp Ile Asn Val Thr Pro Asp Lys Arg Gln Ile Leu Leu						
119	330		335		340		345
121	caa gag gaa aag ctt ttg ttg gca gtt tta aag acc tct ttg ata gga					1107	
122	Gln Glu Glu Lys Leu Leu Leu Ala Val Leu Lys Thr Ser Leu Ile Gly						
123		350		355		360	
125	atg ttt gat agt gat gtc aac aag cta aat gtc agt cag cag cca ctg					1155	
126	Met Phe Asp Ser Asp Val Asn Lys Leu Asn Val Ser Gln Gln Pro Leu						
127		365		370		375	

RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/09/558,149A

TIME: 11:52:44

Input Set : A:\980232_1.TXT

Output Set: N:\CRF4\03142006\I558149A.raw

```

129 ctg gat gtt gaa ggt aac tta ata aaa atg cat gca gcg gat ttg gaa 1203
130 Leu Asp Val Glu Gly Asn Leu Ile Lys Met His Ala Ala Asp Leu Glu
131      380      385      390
133 aag ccc atg gta gaa aag cag gat caa tcc cct tca tta agg act gga 1251
134 Lys Pro Met Val Glu Lys Gln Asp Gln Ser Pro Ser Leu Arg Thr Gly
135      395      400      405
137 gaa gaa aaa aaa gac gtg tcc att tcc aga ctg cga gag gcc ttt tct 1299
138 Glu Glu Lys Lys Asp Val Ser Ile Ser Arg Leu Arg Glu Ala Phe Ser
139 410      415      420      425
141 ctt cgt cac aca aca gag aac aag cct cac agc cca aag act cca gaa 1347
142 Leu Arg His Thr Thr Glu Asn Lys Pro His Ser Pro Lys Thr Pro Glu
143      430      435      440
145 cca aga agg agc cct cta gga cag aaa agg ggt atg ctg tct tct agc 1395
146 Pro Arg Arg Ser Pro Leu Gly Gln Lys Arg Gly Met Leu Ser Ser Ser
147      445      450      455
149 act tca ggt gcc atc tct gac aaa ggc gtc ctg aga cct cag aaa gag 1443
150 Thr Ser Gly Ala Ile Ser Asp Lys Gly Val Leu Arg Pro Gln Lys Glu
151      460      465      470
153 gca gtg agt tcc agt cac gga ccc agt gac cct acg gac aga gcg gag 1491
154 Ala Val Ser Ser Ser His Gly Pro Ser Asp Pro Thr Asp Arg Ala Glu
155      475      480      485
157 gtg gag aag gac tcg ggg cac ggc agc act tcc gtg gat tct gag ggg 1539
158 Val Glu Lys Asp Ser Gly His Gly Ser Thr Ser Val Asp Ser Glu Gly
159 490      495      500      505
161 ttc agc atc cca gac acg ggc agt cac tgc agc agc gag tat gcg gcc 1587
162 Phe Ser Ile Pro Asp Thr Gly Ser His Cys Ser Ser Glu Tyr Ala Ala
163      510      515      520
165 agc tcc cca ggg gac agg ggc tcg cag gaa cat gtg gac tct cag gag 1635
166 Ser Ser Pro Gly Asp Arg Gly Ser Gln Glu His Val Asp Ser Gln Glu
167      525      530      535
169 aaa gcg cct gaa act gac gac tct ttt tca gat gtg gac tgc cat tca 1683
170 Lys Ala Pro Glu Thr Asp Asp Ser Phe Ser Asp Val Asp Cys His Ser
171      540      545      550
173 aac cag gaa gat acc gga tgt aaa ttt cga gtt ttg cct cag cca act 1731
174 Asn Gln Glu Asp Thr Gly Cys Lys Phe Arg Val Leu Pro Gln Pro Thr
175      555      560      565
177 aat ctc gca acc cca aac aca aag cgt ttt aaa aaa gaa gaa att ctt 1779
178 Asn Leu Ala Thr Pro Asn Thr Lys Arg Phe Lys Lys Glu Glu Ile Leu
179 570      575      580      585
181 tcc agt tct gac att tgt caa aag tta gta aat act cag gac atg tca 1827
182 Ser Ser Ser Asp Ile Cys Gln Lys Leu Val Asn Thr Gln Asp Met Ser
183      590      595      600
185 gcc tct cag gtt gat gta gct gtg aaa att aat aag aaa gtt gtg ccc 1875
186 Ala Ser Gln Val Asp Val Ala Val Lys Ile Asn Lys Lys Val Val Pro
187      605      610      615
189 ctg gac ttt tct atg agt tct tta gct aaa cga ata aag cag tta cat 1923
190 Leu Asp Phe Ser Met Ser Ser Leu Ala Lys Arg Ile Lys Gln Leu His
191      620      625      630
193 cat gaa gca cag caa agt gaa ggg gaa cag aat tac agg aag ttt agg 1971

```

RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/09/558,149A

TIME: 11:52:44

Input Set : A:\980232_1.TXT

Output Set: N:\CRF4\03142006\I558149A.raw

```

194 His Glu Ala Gln Gln Ser Glu Gly Glu Gln Asn Tyr Arg Lys Phe Arg
195      635                      640                      645
197 gca aag att tgt cct gga gaa aat caa gca gcc gaa gat gaa cta aga 2019
198 Ala Lys Ile Cys Pro Gly Glu Asn Gln Ala Ala Glu Asp Glu Leu Arg
199 650                      655                      660                      665
201 aaa gag ata agt aaa acg atg ttt gca gaa atg gaa atc att ggt cag 2067
202 Lys Glu Ile Ser Lys Thr Met Phe Ala Glu Met Glu Ile Ile Gly Gln
203                      670                      675                      680
205 ttt aac ctg gga ttt ata ata acc aaa ctg aat gag gat atc ttc ata 2115
206 Phe Asn Leu Gly Phe Ile Ile Thr Lys Leu Asn Glu Asp Ile Phe Ile
207                      685                      690                      695
209 gtg gac cag cat gcc acg gac gag aag tat aac ttc gag atg ctg cag 2163
210 Val Asp Gln His Ala Thr Asp Glu Lys Tyr Asn Phe Glu Met Leu Gln
211                      700                      705                      710
213 cag cac acc gtg ctc cag ggg cag agg ctc ata gca cct cag act ctc 2211
214 Gln His Thr Val Leu Gln Gly Gln Arg Leu Ile Ala Pro Gln Thr Leu
215                      715                      720                      725
217 aac tta act gct gtt aat gaa gct gtt ctg ata gaa aat ctg gaa ata 2259
218 Asn Leu Thr Ala Val Asn Glu Ala Val Leu Ile Glu Asn Leu Glu Ile
219 730                      735                      740                      745
221 ttt aga aag aat ggc ttt gat ttt gtt atc gat gaa aat gct cca gtc 2307
222 Phe Arg Lys Asn Gly Phe Asp Phe Val Ile Asp Glu Asn Ala Pro Val
223                      750                      755                      760
225 act gaa agg gct aaa ctg att tcc ttg cca act agt aaa aac tgg acc 2355
226 Thr Glu Arg Ala Lys Leu Ile Ser Leu Pro Thr Ser Lys Asn Trp Thr
227                      765                      770                      775
229 ttc gga ccc cag gac gtc gat gaa ctg atc ttc atg ctg agc gac agc 2403
230 Phe Gly Pro Gln Asp Val Asp Glu Leu Ile Phe Met Leu Ser Asp Ser
231                      780                      785                      790
233 cct ggg gtc atg tgc cgg cct tcc cga gtc aag cag atg ttt gcc tcc 2451
234 Pro Gly Val Met Cys Arg Pro Ser Arg Val Lys Gln Met Phe Ala Ser
235                      795                      800                      805
237 aga gcc tgc cgg aag tcg gtg atg att ggg act gct ctt aac aca agc 2499
238 Arg Ala Cys Arg Lys Ser Val Met Ile Gly Thr Ala Leu Asn Thr Ser
239 810                      815                      820                      825
241 gag atg aag aaa ctg atc acc cac atg ggg gag atg gac cac ccc tgg 2547
242 Glu Met Lys Lys Leu Ile Thr His Met Gly Glu Met Asp His Pro Trp
243                      830                      835                      840
245 aac tgt ccc cat gga agg cca acc atg aga cac atc gcc aac ctg ggt 2595
246 Asn Cys Pro His Gly Arg Pro Thr Met Arg His Ile Ala Asn Leu Gly
247                      845                      850                      855
249 gtc att tct cag aac tgaccgtagt cactgtatgg aataattggg tttatcgag 2650
250 Val Ile Ser Gln Asn
251                      860
253 atttttatgt tttgaaagac agagtcttca ctaacctttt ttgttttaaa atgaaacctg 2710
254 ctacttaaaaa aaaatacaca tcacacccat ttaaaagtga tcttgagaac cttttcaaac 2770
255 c 2771
257 <210> SEQ ID NO: 2
258 <211> LENGTH: 862

```

RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/09/558,149A

TIME: 11:52:44

Input Set : A:\980232_1.TXT

Output Set: N:\CRF4\03142006\I558149A.raw

```

259 <212> TYPE: PRT
260 <213> ORGANISM: Homo sapiens
262 <400> SEQUENCE: 2
263 Met Glu Arg Ala Glu Ser Ser Ser Thr Glu Pro Ala Lys Ala Ile Lys
264 1 5 10 15
265 Pro Ile Asp Arg Lys Ser Val His Gln Ile Cys Ser Gly Gln Val Val
266 20 25 30
267 Leu Ser Leu Ser Thr Ala Val Lys Glu Leu Val Glu Asn Ser Leu Asp
268 35 40 45
269 Ala Gly Ala Thr Asn Ile Asp Leu Lys Leu Lys Asp Tyr Gly Val Asp
270 50 55 60
271 Leu Ile Glu Val Ser Asp Asn Gly Cys Gly Val Glu Glu Glu Asn Phe
272 65 70 75 80
273 Glu Gly Leu Thr Leu Lys His His Thr Ser Lys Ile Gln Glu Phe Ala
274 85 90 95
275 Asp Leu Thr Gln Val Glu Thr Phe Gly Phe Arg Gly Glu Ala Leu Ser
276 100 105 110
277 Ser Leu Cys Ala Leu Ser Asp Val Thr Ile Ser Thr Cys His Ala Ser
278 115 120 125
279 Ala Lys Val Gly Thr Arg Leu Met Phe Asp His Asn Gly Lys Ile Ile
280 130 135 140
281 Gln Lys Thr Pro Tyr Pro Arg Pro Arg Gly Thr Thr Val Ser Val Gln
282 145 150 155 160
283 Gln Leu Phe Ser Thr Leu Pro Val Arg His Lys Glu Phe Gln Arg Asn
284 165 170 175
285 Ile Lys Lys Glu Tyr Ala Lys Met Val Gln Val Leu His Ala Tyr Cys
286 180 185 190
287 Ile Ile Ser Ala Gly Ile Arg Val Ser Cys Thr Asn Gln Leu Gly Gln
288 195 200 205
289 Gly Lys Arg Gln Pro Val Val Cys Thr Gly Gly Ser Pro Ser Ile Lys
290 210 215 220
291 Glu Asn Ile Gly Ser Val Phe Gly Gln Lys Gln Leu Gln Ser Leu Ile
292 225 230 235 240
293 Pro Phe Val Gln Leu Pro Pro Ser Asp Ser Val Cys Glu Glu Tyr Gly
294 245 250 255
295 Leu Ser Cys Ser Asp Ala Leu His Asn Leu Phe Tyr Ile Ser Gly Phe
296 260 265 270
297 Ile Ser Gln Cys Thr His Gly Val Gly Arg Ser Ser Thr Asp Arg Gln
298 275 280 285
299 Phe Phe Phe Ile Asn Arg Arg Pro Cys Asp Pro Ala Lys Val Cys Arg
300 290 295 300
301 Leu Val Asn Glu Val Tyr His Met Tyr Asn Arg His Gln Tyr Pro Phe
302 305 310 315 320
303 Val Val Leu Asn Ile Ser Val Asp Ser Glu Cys Val Asp Ile Asn Val
304 325 330 335
305 Thr Pro Asp Lys Arg Gln Ile Leu Leu Gln Glu Glu Lys Leu Leu Leu
306 340 345 350
307 Ala Val Leu Lys Thr Ser Leu Ile Gly Met Phe Asp Ser Asp Val Asn
308 355 360 365

```

VERIFICATION SUMMARY

DATE: 03/14/2006

PATENT APPLICATION: US/09/558,149A

TIME: 11:52:45

Input Set : A:\980232_1.TXT

Output Set: N:\CRF4\03142006\I558149A.raw